ARL-TR-8354 ● Apr 2018

US Army Research Laboratory

Machine Learning Intermolecular Potentials for 1,3,5-Triamino-2,4,6-trinitrobenzene (TATB) Using Symmetry-Adapted Perturbation Theory

by DeCarlos E Taylor

Approved for public release; distribution is unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TR-8354 ● APR 2018

US Army Research Laboratory

Machine Learning Intermolecular Potentials for 1,3,5-Triamino-2,4,6-trinitrobenzene (TATB) Using Symmetry-Adapted Perturbation Theory

by DeCarlos E Taylor
Weapons and Materials Research Directorate, ARL

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
April 2018	Technical Report	1 October 2017 – 30 September 2018

4. TITLE AND SUBTITLE
Machine Learning Intermolecular Potentials for 1,3,5-Triamino-2,4,6-trinitrobenzene (TATB) Using Symmetry-Adapted Perturbation Theory

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
DeCarlos E Taylor

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
US Army Research Laboratory
ATTN: RDRL-WML-B
Aberdeen Proving Ground, MD 21005-5069

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-8354

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In this report, intermolecular potentials for 1,3,5-triamino-2,4,6-trinitrobenzene (TATB) are developed using machine learning techniques. Three potentials based on support vector regression, kernel ridge regression, and a neural network are fit using symmetry-adapted perturbation theory. The potentials are used to explore minima on the TATB dimer potential energy surface. It is demonstrated that the ab initio potential energy surface is accurately characterized by the machine learning potentials and that machine learning methods can accurately describe noncovalent interactions in energetic materials.

15. SUBJECT TERMS
machine learning, energetic materials, quantum mechanics, potentials, neural network

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			DeCarlos Taylor
Unclassified	Unclassified	Unclassified	UU	22	19b. TELEPHONE NUMBER (Include area code) 410-306-0853

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

Acknowledgments

This work was supported in part by a grant of computer time from the Department of Defense High Performance Computing Modernization Program at the Army Research Laboratory Department of Defense Supercomputing Resource Center.

INTENTIONALLY LEFT BLANK.

1. Introduction

Recently, there has been a tremendous boom in the application of descriptive and inferential statistical techniques (more popularly known as "machine learning") in a number of fields. The medical industry is using machine learning (ML) and healthcare analytics to assist physicians in the clinical decision-making process.[1] Materials scientists are using ML to design molecules that have specific properties and function.[2] From speech recognition,[3] fraud prevention,[4] spam email filtering,[5] unmanned vehicle operation,[6] finance,[7] and even drunk driver detection,[8] ML has now become an indispensable tool impacting multiple fields and industries.

Computational chemistry is also reaping benefits from ML and it has been used to develop parameters for semi-empirical quantum mechanical (QM) Hamiltonians[9] and for interpolation of ab initio potential energy surfaces.[10] The latter application is particularly appealing as it provides a route toward rapid evaluation of configurational energies and forces using statistical methods that potentially have a QM level of accuracy. Further, since most ML methods only rely on the underlying training data for prediction (hyperparameters and assumptions of statistical distributions aside), they may not suffer the same maladies that plague conventional functional forms such as Tersoff bond order potentials that fail at high pressure due to discontinuous changes in the bond-order term.[11]

ML representations of QM potential energy surfaces have been developed using a variety of ab initio methods including density functional theory (DFT) and coupled cluster theory for covalently bonded systems. However, for energetic molecular crystals of interest to the Army, accurate representation of noncovalent interactions is critical. Gao et al.[12] used ML to develop van der Waals corrections for DFT. McGibbon et al.[13] developed a hybrid ML and QM methodology for computation of interaction energies where a neural network, trained using coupled cluster reference data, was used to correct Moller-Plesset (MP2) interaction energies. Using this combined approach, they obtained a 6-fold improvement in accuracy relative to conventional MP2.

Given the importance of noncovalent interactions in energetic molecular crystals, Army scientists have focused on development and application of QM approaches that can accurately describe weak electron correlations ("dispersion") between molecules. One successful technique uses a combination of DFT and symmetry-adapted perturbation theory (SAPT) known as SAPT(DFT).[14] In SAPT(DFT), the intramonomer correlation is treated through the exchange-correlation density functional yielding a single perturbative series representing the intermolecular interaction which, when combined with resolution of the identity techniques,

reduces the computational scaling. Accurate dispersion energies, of particular importance for the interlayer interactions in molecular crystals, are obtained via application of the generalized Casimir-Polder expression with frequency dependent density susceptibilities computed within the coupled Kohn-Sham approach. Given its combination of accuracy and efficiency, SAPT(DFT) has been used to develop intermolecular potentials for the cyclotrimethylene trinitramine,[15] 1,1-diamino-2,2-dinitroethylene,[16] and 1,3,5-triamino-2,4,6-trinitrobenzene[17] (TATB) energetics using conventional functional forms such as exponential-6 (Exp-6).

In this work, the previously computed SAPT(DFT) reference data for TATB were used to develop new intermolecular potentials using 3 ML methods:

- Support vector regression[18]
- Kernel ridge regression[19]
- Neural networks[20]

These ML models differ from the previous work[17] that used a parametric function of Exp-6 form to fit the SAPT(DFT) reference data. The Exp-6 potential requires 168 descriptors for each dimer configuration (144 Cartesian coordinates and 24 charges) as input. On the contrary, the new ML models use a reduced descriptor set and only require 6 input variables: the center of mass separation between monomers (R) and 5 Euler angles defining the monomer orientations. The ML models were applied to potential energy surface cross sections of minima reported in the previous work.[17] It is observed that stable dimer configurations are accurately described by the ML models and that support vector regression (SVR), kernel ridge regression (KRR), and neural networks can be used to accurately interpolate SAPT(DFT) surfaces.

2. Computational Methods

2.1 Quantum Mechanical Reference Data

The ML methods were trained using a grid of 880 randomly configured TATB dimer configurations computed previously using SAPT(DFT). Full details of the SAPT(DFT) calculations are given in Taylor[17]. In summary, the geometry of the TATB monomer (Fig. 1) used in the calculations was taken from the experimental unit cell and all interaction energy calculations used an aug-cc-pVDZ basis supplemented by a set of 3s(α = 0.9,0.3,0.1) 3p(α = 0.9,0.3,0.1) 2d(α = 0.6,0.2) 2f(α = 0.6,0.2) “midbond” functions with a PBE0 density functional description of the monomers. In the previous work,[17] the SAPT(DFT) interaction energies were fit to an Exp-6 functional form using the PIKAIA[21] genetic algorithm with a

population of 100 individuals that evolved for 500 generations. Fitness scoring of the individuals in the population was given by the magnitude of the root mean square deviation between the fitted and reference interaction energies.

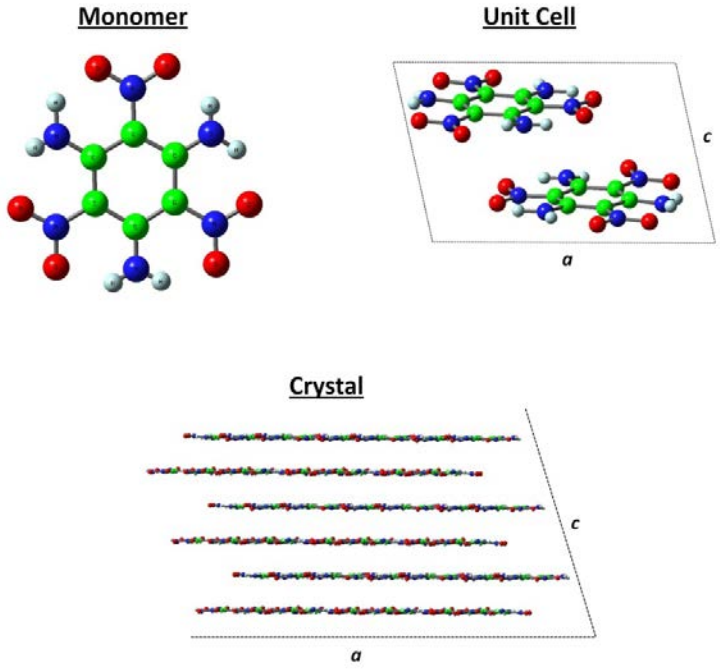

Fig. 1 Molecular and condensed phase structure of TATB. Carbons are green, nitrogens are blue, oxygens are red, and hydrogens are white.

2.2 Machine Learning Methods

Using the SAPT(DFT) reference energies, 3 ML potentials were developed using SVR, KRR, and a neural network of radial basis functions. The SVR and KRR potentials were developed using the sklearn[22] modules available in Python and the neural network was implemented from scratch using Python code developed by the author.

2.2.1 Support Vector Regression

SVR is an extension of the support vector classifier and is used for prediction of quantitative instead of categorical variables.[18] In general, given a set of predictor variables *x*, one wishes to optimize the weights *w*, and intercept *b*, of the following function:

$$y(x) = w\emptyset(x) + b\ , \qquad (1)$$

where $\emptyset(x)$ is a transformation of the feature space. The weights and intercept are obtained by minimizing the constrained objective function:

$$\min_{w}\left\{\frac{1}{2}\|w\|^2 + C\sum_i(\xi_i + \xi_i^*)\right\}$$

$$s.t.\begin{cases} y_i - w\emptyset(x) - b \le \varepsilon + \xi_i \\ -y_i + w\emptyset(x) + b \le \varepsilon + \xi_i^* \\ \xi_i \ge 0\ ;\ \xi_i^* \ge 0 \end{cases}, \tag{2}$$

where *C*, $\boldsymbol{\varepsilon}$, and $\boldsymbol{\xi}$ collectively control the maximum allowable error. The feature space can be expanded by reformulating the above expressions in terms of dot products of the predictors *x*. When written in terms of dot products, one can then take advantage of the "kernel trick", which allows sampling of transformed feature spaces of increased dimension (infinite for some kernels) without having to explicitly sample the larger space. In this work, the radial basis function kernel was used:

$$K(x_i, x) = e^{-\gamma\|x_i - x\|^2}\ , \tag{3}$$

and the parameters C and γ in Eqs. 2 and 3, respectively, were determined using cross-validation (discussed in Section 2.2.4).

2.2.2 Kernel Ridge Regression

KRR can be derived from Eq. 2 by ignoring the bias term b, setting ε = 0, and squaring the "slack variables" ξ. This yields the following expression:

$$\min_{w}\{a\|w\|^2 + \sum_i[y_i - w\emptyset(x)]^2\}\ , \tag{4}$$

where *a* serves as a regularization parameter. KRR is also amenable to the kernel trick, and the radial basis function kernel of Eq. 3 was also used with hyperparameters determined by cross-validation.

2.2.3 Neural Network

Neural networks generally consist of input and output layers that are separated by hidden layers with nodes adjoined by weighted connections (Fig. 2). The weights can be determined by minimization of a loss function (squared error for example) using back propagation[20] for analytically differentiable activation functions or with stochastic optimization methods such as genetic algorithms.[23] In this work, a feed-forward network was developed using one hidden layer of radial basis functions, centered on the SAPT(DFT) reference configurations, with a vector containing the center of mass separation and Euler angles fed into the input layer. When using a

single input node and one hidden layer, the weights of the network can be solved analytically by inversion of the Gram matrix K(x, x') that is computed using the radial basis function kernel in Eq. 3. The exponent γ defining the Gram matrix was determined by cross-validation and all basis functions used the same exponent.

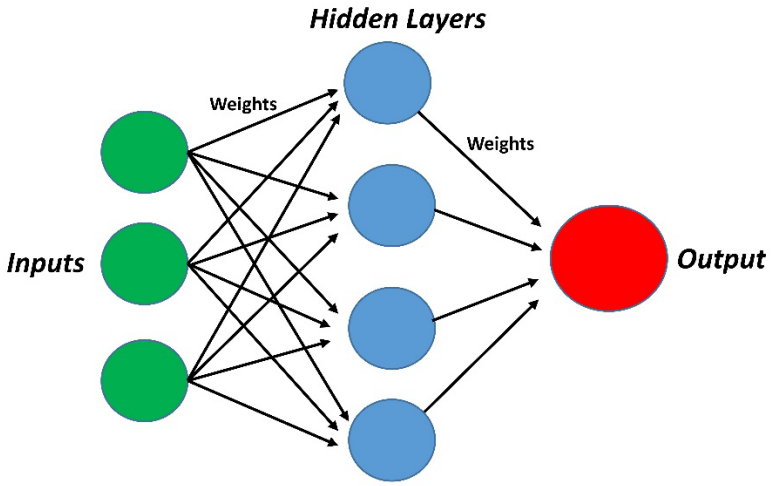

Fig. 2 General schematic of a neural network showing input nodes, hidden layer, and output node with weighted connections

2.2.4 Cross Validation

The hyperparameters for each model were determined by splitting the SAPT(DFT) reference data into a training set (766 configurations) and test set (86 configurations) followed by a grid search over hyperparameters to maximize the coefficient of determination (Q^2) for the test set. Correlation plots for the test set using each ML method are shown in Fig. 3 and the best agreement for this test set was obtained using KRR with $Q^2 = 0.85$. The out-of-sample performance for each ML potential is deemed acceptable given the paucity of geometric information used as input for each configuration. The fits could likely be improved by using the full set of Cartesian coordinates for each configuration, thereby increasing the dimension of the predictor space. However, given that SAPT(DFT) is a rigid monomer, intermolecular theory, the internal coordinates would be somewhat extraneous in the current context.

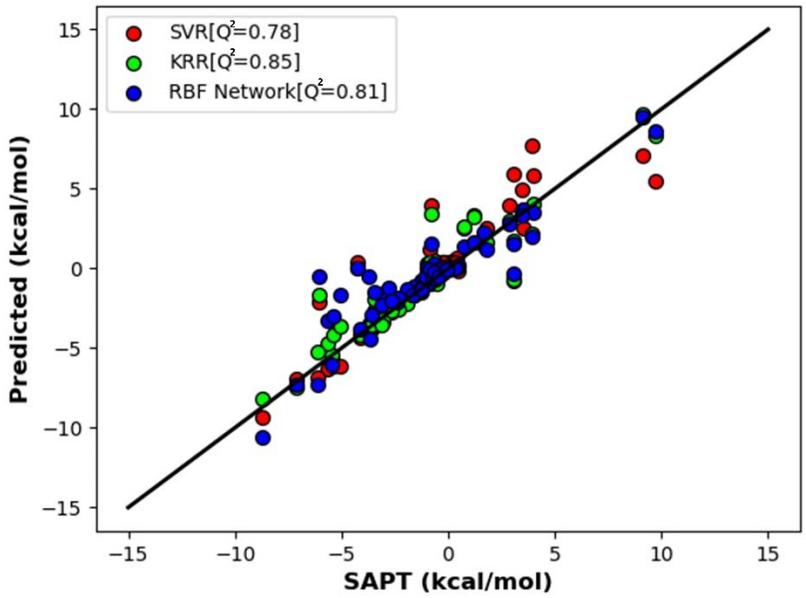

Fig. 3 Interaction energy correlation plots for SVR, KRR, and neural network potentials on test set with 86 configurations

3. Potential Energy Surface Characterization

The ML potentials were used to compute potential energy surface cross sections of 8 stable configurations of TATB dimers previously identified in Taylor.[17] The structures and cross sections using the ML potentials, the Exp-6 potential from Taylor,[17] and the SAPT(DFT) energies are presented in Figs. 4 and 5. As shown, the topology of the ML potential energy surfaces are in good agreement with the ab initio data and stable minima on the ab initio surface are also present on the ML surfaces. Clearly one cannot locate all minima on the surface nor can it be guaranteed that all minima on the fitted surfaces correspond to minima on the ab initio surface. However, for the configurations presented in Taylor,[17] the ab initio potential energy surface is well characterized by the ML potentials.

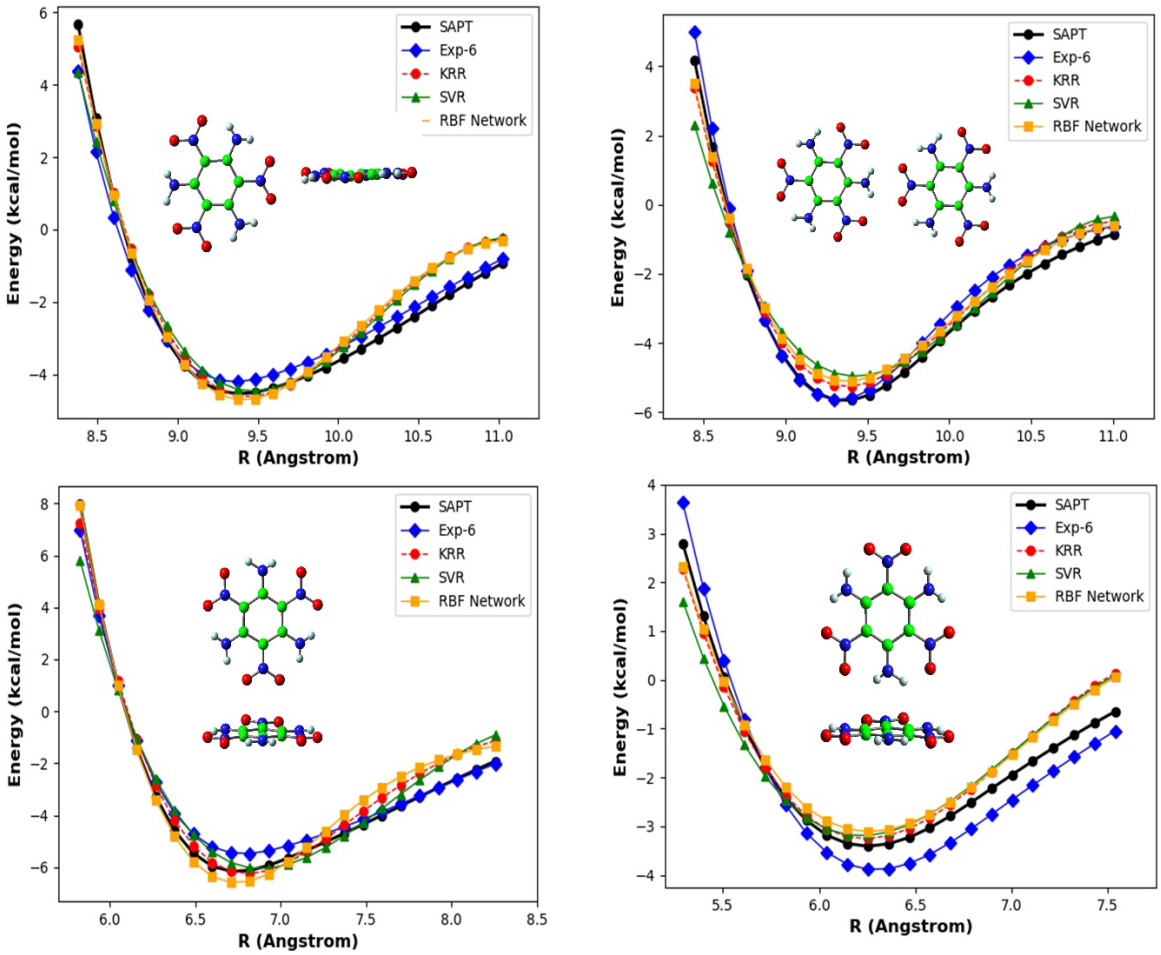

Fig. 4 Potential energy surface cross sections for nitro and amine substituent interactions

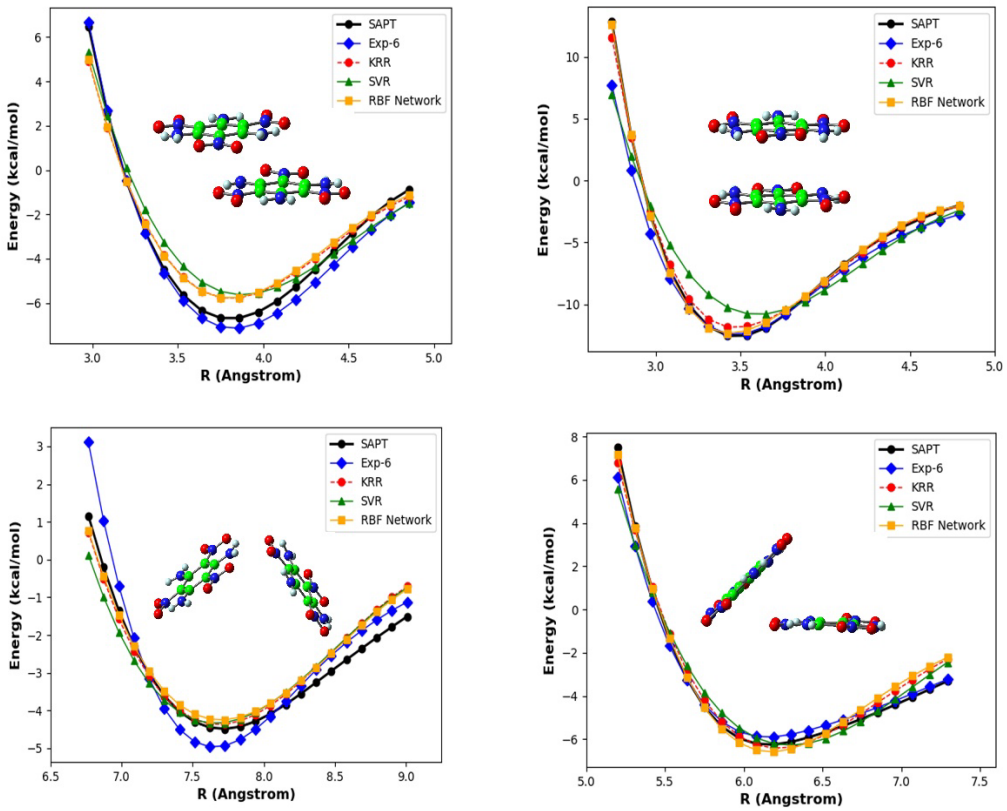

Fig. 5 Potential energy surface cross sections for ring interactions

4. Conclusion

Although the ML potentials perform well for the stationary points presented in this work, as with any fitted model, there will be regimes where the accuracy is less than optimal. As an example, during testing it was observed that at a large intermonomer separation (R>12 angstrom) for some sample configurations, the SVR potential had a non-zero (but still small) interaction energy of approximately –0.1 kcal/mol. This is to be compared to the KRR and RBF potentials that predicted energies of magnitude less than 10^{-10} kcal/mol, as one would generally expect at large separation. This is likely due to a diffuse exponent used in the radial basis function kernel of the SVR potential that results in non-negligible contributions to the interaction energy, even at large separation. In practice, this could be remedied by including more asymptotic points in the training set to obtain a better exponent for the kernel or by simply using a cutoff for the potential.

It is possible to re-fit the ML potentials using Cartesian coordinates for the dimers instead of the reduced set of descriptors used in this work. Cartesian potentials, and the associated forces, could then be used to perform molecular dynamics

simulations to determine if the ML potentials can accurately reproduce the condensed phase crystal structure. This work is underway and will be the subject of a future report.

5. References

1. Bennett CC, Hauser K. Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. Art Int in Med. 2013;57(1):9–19.

2. Gomez-Bombarelli R, Aguilera-Iparraguirre J, Hirzel TD, Duvenaud D, Maclaurin D, Blood-Forsythe MA, Chae HS, Einzinger M, Ha DG, Wu T, et al. Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. Nat Mat. 2016;15(10):1120–1127.

3. Saon G, Picheny M. Recent advances in conversational speech recognition using conventional and recurrent neural networks. IBM J Res Dev. 2017;61(4):1–10.

4. Triepels R, Daniels H, Feelders A. Data-driven fraud detection in international shipping. Exp Sys Appl. 2018;99:193–202.

5. Shams R, Mercer R. Supervised classification of spam emails with natural language stylometry. Neu Comp App. 2016;27(8):2315–2331.

6. Ammour N, Alhichri H, Bazi Y, Benjdira B, Alajlan N, Zuair M. Deep learning approach for car detection in UAV imagery. Rem Sens. 2017;9(4):312–327.

7. Heaton J, Polson N, Witte J. Deep learning for finance: deep portfolios. App Sto Mod Bus Ind. 2017;33(1):3–12.

8. Chen H, Chen L. Support vector machine classification of drunk driving behaviour. Int J Env Res Public Health. 2017;14(1):1–14. Article 108. Doi:10.3390/ijerph14010108.

9. Dral P, von Lilienfeld A, Thiel W. Machine learning of parameters for accurate semiempirical quantum chemical calculations. J Chem Theo Comp. 2015;11(5):2120–2125.

10. Zuo J, Zhao B, Guo H, Xie D. A global coupled cluster potential energy surface for HCl + OH ↔ Cl + H_2O. Phy Chem Chem Phys. 2017;19:9770–9777.

11. Porter L, Yip S, Yamaguchi M, Kaburaki H, Tang M. Empirical bond-order potential description of thermodynamic properties of crystalline silicon. J Appl Phy. 1997;8(1):96–106.

12. Gao T, Li H, Li W, Li L, Fang C, Li H, Hu L, Lu Y, Su Z-M. A machine learning correction for DFT non-covalent interactions on the S22, S66 and X40 benchmark databases. J Cheminf. 2016;8(24):1–17.

13. McGibbon R, Taube1 AG, Donchev AG, Siva K, Hernández F, Hargus C, Law K-H, Klepeis JL, and Shaw DE. Improving the accuracy of Moller-Plesset perturbation theory with neural networks. J Chem Phys. 2017;147:161725(1–15).

14. Williams H, Chabalowski C. Using Kohn-Sham orbitals in symmetry-adapted perturbation theory to investigate intermolecular interations. J Phys Chem A. 2001;105(3):646–659.

15. Podeszwa R, Bukowski R, Rice B, Szalewicz K. Potential energy surface for cyclotrimethylenetrinitramine dimer from symmetry-adapted perturbation theory. Phys Chem Chem Phys. 2007;9(41):5561–5569.

16. Taylor D, Rob F, Rice B, Podeszwa R, Szalewicz K. A molecular dynamics study of 1,1-diamino-2,2-dinitroethylene (FOX-7) crystal using a symmetry adapted perturbation theory-based intermolecular force field. Phys Chem Chem Phys. 2011;13(37):16629–16636.

17. Taylor D. Intermolecular forces and molecular dynamics simulation of 1,3,5-triamino-2,4,6-trinitro (TATB) using symmetry adapted perturbation theory. J Phys Chem A. 2013;117(16):3507–3520.

18. Basak D, Pal S, Patranabis D. Support vector regression. Neur Inf Proc Lett. 2007;11(10):203–224.

19. Vovk V. Kernel ridge regression. In: Empirical inference. Berlin (Germany): Springer; 2013. p. 105–116.

20. Krisel D. A brief introduction to neural networks. 2007. www.drkriesel.com.

21. Charbonneau P. Genetic algorithms in astronomy and astrophysics. Astrophy Jou. 1995;101:309.

22. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. Scikit-learn: machine learning in python. J Mach Lea Res. 2011;12:2825–2830.

23. Maniezzo V. Genetic evolution of the topology and weight distribution of neural networks. IEEE T Neur Net. 1994;5(1):39–53.

Approved for public release; distribution is unlimited.

List of Symbols, Abbreviations, and Acronyms

DFT	density functional theory
Exp-6	exponential-6
KRR	kernel ridge regression
ML	machine learning
MP2	Moller-Plesset
QM	quantum mechanical
SAPT	symmetry-adapted perturbation theory
SVR	support vector regression
TATB	1,3,5-triamino-2,4,6-trinitrobenzene

1 (PDF) DEFENSE TECHNICAL INFORMATION CTR DTIC OCA

2 (PDF) DIR ARL
IMAL HRA
RECORDS MGMT
RDRL DCL
TECH LIB

1 (PDF) GOVT PRINTG OFC
A MALHOTRA

1 (PDF) ARMY RESEARCH OFC
J PARKER

19 (PDF) ARL
RDRL D
M TSCHOPP
RDRL WM
B FORCH
RDRL WML B
N TRIVEDI
B BARNES
I BATYREV
J BRENNAN
E BYRD
J LARENTZOS
W MATTSON
B RICE
D TAYLOR
N WEINGARTEN
RDRL WML D
M MCQUAID
J VEALS
RDRL WMM B
E HERNANDEZ
RDRL WMM F
S COLEMAN
RDRL WMM G
J ANDZELM
B RINDERSPACHER
RDRL SER E
M NEUPANE

INTENTIONALLY LEFT BLANK.